

Numerical methods for a Kohn-Sham density functional model based on optimal transport

Huajie Chen ^{*}, Gero Friesecke [†] and Christian B. Mendl [‡]

May 28, 2014

Abstract

In this paper, we study numerical discretizations to solve density functional models in the “strictly correlated electrons” (SCE) framework. Unlike previous studies our work is not restricted to radially symmetric densities. In the SCE framework, the exchange-correlation functional encodes the effects of the strong correlation regime by minimizing the pairwise Coulomb repulsion, resulting in an optimal transport problem. We give a mathematical derivation of the self-consistent Kohn-Sham-SCE equations, construct an efficient numerical discretization for this type of problem for $N = 2$ electrons, and apply it to the H_2 molecule in its dissociating limit. Moreover, we prove that the SCE density functional model is correct for the H_2 molecule in its dissociating limit.

1 Introduction

In the ab-initio quantum mechanical modeling of many-particle systems, Kohn-Sham density functional theory (DFT) [1, 2] achieves so far the best compromise between accuracy and computational cost, and has become the most widely used electronic structure model in molecular simulations and material science. Within the traditional Kohn-Sham formulation, the ground state energy and electron density of an N -electron system can be obtained by minimizing the energy functional

$$E_{\text{KS}}[\{\phi_i\}_{i=1}^N] = \int_{\mathbb{R}^3} \left(\frac{1}{2} \sum_{i=1}^N |\nabla \phi_i(\mathbf{r})|^2 + v_{\text{ext}}(\mathbf{r}) \rho(\mathbf{r}) \right) d\mathbf{r} + E_{\text{H}}[\rho] + E_{\text{xc}}[\rho] \quad (1.1)$$

with respect to orbitals $\{\phi_i\}_{i=1}^N$ under the constraint $\int_{\mathbb{R}^3} \phi_i \phi_j = \delta_{ij}$. Here, $\rho(\mathbf{r}) = \sum_{i=1}^N |\phi_i(\mathbf{r})|^2$ is the electron density, v_{ext} is the electrostatic attraction potential generated by the nuclei, $E_{\text{H}}[\rho] = \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}'$ is the Hartree energy that describes electron-electron Coulomb repulsion energy by

^{*}Zentrum Mathematik, Technische Universität München, Boltzmannstraße 3, 85747 Garching, Germany. E-mail: chenh@ma.tum.de.

[†]Zentrum Mathematik, Technische Universität München, Boltzmannstraße 3, 85747 Garching, Germany. E-mail: gf@ma.tum.de.

[‡]Zentrum Mathematik, Technische Universität München, Boltzmannstraße 3, 85747 Garching, Germany. E-mail: mendl@ma.tum.de.

a mean field approximation, and $E_{\text{xc}}[\rho]$ is the so-called exchange-correlation energy functional that includes all the many-particle interactions.

The major drawback of DFT is the fact that the exact functional for the exchange-correlation energy is not known. A basic model is local density approximation (LDA) [2, 3], which is still commonly used in practical calculations. Improvements of this model give rise to the generalized gradient approximation (GGA) [4, 5, 6] and hybrid functionals [7, 8, 9]. Although these models have achieved high accuracy for many chemical and physical systems, there remain well-known limitations. For example, in systems with significant static correlation [10], LDA, GGA, and also hybrid functionals underestimate the magnitude of the correlation energy. This becomes particularly problematic for the dissociation of electron pair bonds. A famous example is the dissociating H_2 molecule: the widely employed LDA, GGA, and even hybrid models fail rather badly at describing the energy curve for dissociating H_2 . Many efforts have been made in order to make an appropriate ansatz for the exchange-correlation functional and tackle this problem (e.g., [11, 12]). In our view, a principal deficiency of these works is the attempt to describe strong correlation within the framework of mean field approximations.

Alternatively, DFT calculations can also be based on the strongly interacting limit of the Hohenberg-Kohn density functional, denoted “strictly correlated electrons” (SCE) DFT [13, 14]. This approach considers a reference system with complete correlation between the electrons, and is able to capture key features of strong correlation within the Kohn-Sham framework. The pioneering work [15, 16, 14] has shown that the SCE ansatz can describe certain model systems in the extreme strongly correlated regime with higher accuracy than standard Kohn-Sham DFT. However, the calculations are presently limited to either one-dimensional or spherically symmetric systems. To our knowledge, there is no SCE-DFT calculation for dissociating the H_2 molecule in \mathbb{R}^3 .

In the SCE-DFT model, the repulsion energy between strongly interacting electrons is related to optimal transport theory. Optimal transport was historically studied in [17] to model the most economical way of moving soil from one area to another, and was further generalized in [18, 19] to the Kantorovich primal and dual formulation. The goal is to transfer masses from an initial density ρ_A to a target density ρ_B in an optimal way such that the “cost” $c(x, y)$ for transporting mass from x to y is minimized (see [20] for a comprehensive treatment). The Coulomb repulsion energy in the SCE-DFT model can be reformulated as the optimal cost of an optimal transport problem, if we identify the marginals with the electron density divided by number of electrons, i.e., ρ/N , and the cost function with the electron-electron Coulomb repulsion

$$c_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{1 \leq i < j \leq N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (1.2)$$

For instance, for a two-electron system within the SCE-DFT framework, the electron repulsion energy for a given single-particle electron density ρ is

$$V_{\text{ee}}^{\text{SCE}}[\rho] = \min_{\Psi} \left\{ \int_{\mathbb{R}^6} \frac{|\Psi(\mathbf{r}_1, \mathbf{r}_2)|^2}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2, \right. \\ \left. \int_{\mathbb{R}^3} |\Psi(\mathbf{r}_1, \mathbf{r})|^2 d\mathbf{r}_1 = \int_{\mathbb{R}^3} |\Psi(\mathbf{r}, \mathbf{r}_2)|^2 d\mathbf{r}_2 = \frac{\rho(\mathbf{r})}{2} \right\}. \quad (1.3)$$

Strictly speaking, the set of admissible $|\Psi|^2$'s must be enlarged to probability measures in order to allow strict correlation, which corresponds to concentration of the many-body probability density on a lower dimensional subset [21]. There are several mathematical investigations of the relations between SCE-DFT and optimal transport problems, see [22, 21, 23, 24], but important open problems remain. To our knowledge, the functional derivative of the SCE functional (1.3) (alias optimal cost functional) with respect to the electron density (alias marginal measure) is not clear from a mathematical point of view. However, this result is crucial for deriving the Kohn-Sham equations needed in practical calculations. Numerical algorithms for optimal transport problems are rather sparse. Explicit solutions for the co-motion functions are known for one-dimensional and spherically symmetric problems [14], but cannot be generalized to two- and three-dimensional systems. An alternative route might be the Kantorovich dual formulation of the SCE functional [22, 25]. In a complementary work [26], the H_2 molecule is studied using an ansatz for the dual potential, and there is a recent simulation of a one-dimensional model H_2 molecule using the SCE framework [27].

In this paper, we give a mathematical derivation of the Kohn-Sham equations for optimal transport-based DFT which is rigorous up to physically expected smoothness and continuity assumptions (section 3), provide an efficient numerical algorithm for discretising and solving the resulting optimal transport problem for the case of two electrons without restriction to radial symmetry (section 4), and then apply this algorithm to a self-consistent DFT simulation of the H_2 molecule in the dissociating limit (section 5). Finally, we show both numerically and by a rigorous mathematical argument that the SCE-DFT model is accurate for the H_2 molecule in the dissociating limit.

2 Preliminaries

Consider a molecular system with M nuclei of charges $\{Z_1, \dots, Z_M\}$, located at positions $\{\mathbf{R}_1, \dots, \mathbf{R}_M\}$, and N electrons in the non-relativistic setting. The electrostatic potential generated by the nuclei is

$$v_{\text{ext}}(\mathbf{r}) = - \sum_{I=1}^M \frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|}, \quad \mathbf{r} \in \mathbb{R}^3.$$

Within the DFT framework [1, 28], the ground state density and energy of the system is obtained by solving the following minimization problem

$$E_0 = \min_{\rho} \left\{ F_{\text{HK}}[\rho] + \int_{\mathbb{R}^3} v_{\text{ext}} \rho, \quad \rho \geq 0, \quad \sqrt{\rho} \in H^1(\mathbb{R}^3), \quad \int_{\mathbb{R}^3} \rho = N \right\}, \quad (2.1)$$

where ρ is the electron density and $F_{\text{HK}}[\rho]$ is the so-called Hohenberg-Kohn functional [1]. F_{HK} is a universal functional of ρ in the sense that it does not depend on the external potential v_{ext} . Unfortunately, no tractable expression for F_{HK} is known that could be used in numerical simulations. The standard Kohn-Sham DFT [2] treats the system as N non-interacting electrons, and ap-

proximates $F_{\text{HK}}[\rho]$ by a summation of the kinetic energy

$$T_{\text{KS}}[\rho] = \inf \left\{ \frac{1}{2} \sum_{i=1}^N \int_{\mathbb{R}^3} |\nabla \phi_i(\mathbf{r})|^2 d\mathbf{r}, \phi_i \in H^1(\mathbb{R}^3), \right. \\ \left. \sum_{i=1}^N |\phi_i(\mathbf{r})|^2 = \rho(\mathbf{r}), \int_{\mathbb{R}^3} \phi_i \phi_j = \delta_{ij} \right\}, \quad (2.2)$$

the Hartree energy $E_{\text{H}}[\rho]$, and an exchange-correction energy $E_{\text{xc}}[\rho]$, as shown in Eq. (1.1).

Since the standard non-interacting model cannot capture the features that result from strong correlation, it is not able to simulate strongly correlated electron systems, like the H_2 molecule in its dissociating limit. In contrast to that, the SCE-DFT model [29, 30, 14] starts from the strongly interacting limit (semi-classical limit) of F_{HK} , and gives rise to the following SCE functional (see [24] for a mathematical justification)

$$V_{\text{ee}}^{\text{SCE}}[\rho] = \inf \{ V_{\text{ee}}[\rho_N], \rho_N(\mathbf{r}_1, \dots, \mathbf{r}_N) \geq 0, \rho_N \text{ is symmetric, } \rho_N \mapsto \rho \}, \quad (2.3)$$

where

$$V_{\text{ee}}[\rho_N] = \int_{\mathbb{R}^{3N}} \sum_{1 \leq i < j \leq N} \frac{\rho_N(\mathbf{r}_1, \dots, \mathbf{r}_N)}{|\mathbf{r}_i - \mathbf{r}_j|} d\mathbf{r}_1 \cdots d\mathbf{r}_N, \quad (2.4)$$

and $\rho_N \mapsto \rho$ means that ρ is the marginal distribution of ρ_N , that is to say

$$\rho(\mathbf{r}) = N \int_{\mathbb{R}^{3(N-1)}} \rho_N(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) d\mathbf{r}_2 \cdots d\mathbf{r}_N.$$

The minimization in Eq. (2.3) is over all symmetric N -point probability measures ρ_N which have the given single-particle density ρ as marginal, and yields the minimum of the electronic Coulomb repulsion energy over all such ρ_N . The SCE-DFT model takes $V_{\text{ee}}^{\text{SCE}}[\rho]$ as the only interaction term, replacing $E_{\text{H}}[\rho] + E_{\text{xc}}[\rho]$ in standard Kohn-Sham DFT.

The minimization task (2.3) is in fact an optimal transport problem with Coulomb cost [21, 22, 24], which has two alternative formulations: the Monge formulation and the Kantorovich dual formulation. For the Monge formulation, one uses the ansatz

$$\rho_N(\mathbf{r}_1, \dots, \mathbf{r}_N) = \frac{\rho(\mathbf{r}_1)}{N} \delta(\mathbf{r}_2 - T_2(\mathbf{r}_1)) \cdots \delta(\mathbf{r}_N - T_N(\mathbf{r}_1)) \quad (2.5)$$

with $T_i : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ($i = 2, \dots, N$) the so-called co-motion functions (also called optimal transport maps), where we use the convention $T_1(\mathbf{r}) = \mathbf{r}$. The above ansatz already appears on physical grounds, without reference to optimal transport theory, in [14]. Since the N -particle distribution ρ_N in (2.5) is zero everywhere except on the set

$$M = \{(\mathbf{r}, T_2(\mathbf{r}), \dots, T_N(\mathbf{r})), \mathbf{r} \in \mathbb{R}^3\}, \quad (2.6)$$

it describes a state where the location of one electron fixes all the other $N - 1$ electrons through the co-motion functions T_i , $i = 2, \dots, N$. The co-motion functions are implicit functionals of the density, determined by the minimization

problem (2.3) and a set of differential equations that ensure the invariance of the density under the coordinate transformation $\mathbf{r} \mapsto T_i(\mathbf{r})$ [14], i.e.,

$$\rho(T_i(\mathbf{r}))dT_i(\mathbf{r}) = \rho(\mathbf{r})d\mathbf{r}. \quad (2.7)$$

In terms of these functions, the optimal value of (2.3) reads

$$V_{\text{ee}}^{\text{SCE}}[\rho] = \frac{1}{N} \int_{\mathbb{R}^3} \sum_{1 \leq i < j \leq N} \frac{\rho(\mathbf{r})}{|T_i(\mathbf{r}) - T_j(\mathbf{r})|} d\mathbf{r}. \quad (2.8)$$

Note that the ansatz (2.5) is not in general symmetric under exchanging particle coordinates, nevertheless, dropping the symmetrization does not alter the minimum value of (2.3).

Alternatively, one can start from the so-called Kantorovich dual formulation [19]. It has been shown in [22] that the value of $V_{\text{ee}}^{\text{SCE}}[\rho]$ is exactly given by the maximum of this Kantorovich dual problem

$$V_{\text{ee}}^{\text{SCE}}[\rho] = \max \left\{ \int_{\mathbb{R}^3} u(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}, \sum_{i=1}^N u(\mathbf{r}_i) \leq \sum_{1 \leq i < j \leq N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \right\}. \quad (2.9)$$

In what follows, we denote the maximizer of (2.9) by u_ρ , which is called the Kantorovich potential. We assume that u_ρ is unique and depends continuously on ρ in the sense that

$$u_{\rho_j} \xrightarrow{*} u_\rho \quad \text{if} \quad \rho_j \rightarrow \rho \text{ in } L^1(\mathbb{R}^3). \quad (2.10)$$

For numerical implementations, the Kantorovich dual formulation has high complexity due to the $3N$ -dimensionality of the constraints. In comparison, the Monge formulation amounts to a spectacular dimension reduction, in which the unknowns are $N - 1$ maps on \mathbb{R}^3 instead of one function ρ_N on \mathbb{R}^{3N} . However, for practical purposes it is currently restricted to spherically symmetric densities and one-dimensional systems, for which the constraints (2.7) can be solved semi-analytically. Our purpose is to construct an efficient numerical discretization of the Monge formulation for $N = 2$ electrons which is applicable to non-spherical systems.

3 Kohn-Sham equations for optimal transport based DFT

By taking $V_{\text{ee}}^{\text{SCE}}$ as the only interaction term within the Kohn-Sham DFT framework, we can obtain the ground state approximations of energy and electron density by solving the following minimization problem

$$E_0 = \inf_{\Phi} \left\{ E_{\text{KS}}^{\text{SCE}}[\Phi], \phi_i \in H^1(\mathbb{R}^3), \int_{\mathbb{R}^3} \phi_i \phi_j = \delta_{ij} \right\}, \quad (3.1)$$

where $\Phi = (\phi_1, \dots, \phi_N)$ denotes the Kohn-Sham orbitals and

$$E_{\text{KS}}^{\text{SCE}}[\Phi] = \frac{1}{2} \sum_{i=1}^N \int_{\mathbb{R}^3} |\nabla \phi_i(\mathbf{r})|^2 d\mathbf{r} + \int_{\mathbb{R}^3} v_{\text{ext}}(\mathbf{r})\rho_{\Phi}(\mathbf{r})d\mathbf{r} + V_{\text{ee}}^{\text{SCE}}[\rho_{\Phi}] \quad (3.2)$$

with $\rho_\Phi(\mathbf{r}) = \sum_{i=1}^N |\phi_i(\mathbf{r})|^2$. We shall derive the self-consistent Kohn-Sham equations for (3.1) in this section. The key point is to calculate the functional derivative of the SCE functional $\delta V_{\text{ee}}^{\text{SCE}}[\rho]/\delta\rho$ with respect to the single particle density ρ , which is the effective one-body potential coming from the interaction term $V_{\text{ee}}^{\text{SCE}}[\rho]$. In the derivation below, we make various plausible assumptions on the Kantorovich potential such as uniqueness, continuous dependence on the density, and differentiability at relevant points. We believe these assumptions to be correct except possibly in exceptional situations. A fully rigorous treatment without these assumptions would be desirable, but lies beyond the scope of this paper.

Note that the functional $V_{\text{ee}}^{\text{SCE}}[\rho]$ is not defined on arbitrary densities, but only on those with $\int_{\mathbb{R}^3} \rho = N$. Therefore, the definition of the functional derivative only specifies its integral against perturbations in the corresponding “tangent space”, that is to say perturbations that have integral zero:

$$\int_{\mathbb{R}^3} \frac{\delta V_{\text{ee}}^{\text{SCE}}[\rho]}{\delta\rho} \cdot \tilde{\rho} = \lim_{\varepsilon \rightarrow 0} \frac{V_{\text{ee}}^{\text{SCE}}[\rho + \varepsilon\tilde{\rho}] - V_{\text{ee}}^{\text{SCE}}[\rho]}{\varepsilon} \quad \text{for all } \tilde{\rho} \text{ with } \int_{\mathbb{R}^3} \tilde{\rho} = 0. \quad (3.3)$$

The following theorem indicates that the functional derivative is nothing but the Kantorovich potential u_ρ with an additive constant.

Theorem 3.1. *Assume that the maximizer u_ρ of (2.9) is unique and depends continuously on the electron density ρ in the sense of (2.10). Then $v_{\text{SCE}}[\rho] = \delta V_{\text{ee}}^{\text{SCE}}[\rho]/\delta\rho$ is a functional derivative of $V_{\text{ee}}^{\text{SCE}}$ at point ρ in the sense of (3.3) if and only if*

$$v_{\text{SCE}}[\rho] = u_\rho + C \quad \text{for any constant } C. \quad (3.4)$$

Proof. For any given single-particle density ρ with $\int_{\mathbb{R}^3} \rho = N$, and any perturbation $\tilde{\rho}$ with $\int_{\mathbb{R}^3} \tilde{\rho} = 0$, we define

$$D_\varepsilon = \frac{V_{\text{ee}}^{\text{SCE}}[\rho + \varepsilon\tilde{\rho}] - V_{\text{ee}}^{\text{SCE}}[\rho]}{\varepsilon}.$$

For simplicity, we assume $\varepsilon > 0$. We have from (2.9) that

$$D_\varepsilon = \frac{\int_{\mathbb{R}^3} u_{\rho+\varepsilon\tilde{\rho}}(\rho + \varepsilon\tilde{\rho}) - \int_{\mathbb{R}^3} u_\rho \rho}{\varepsilon}. \quad (3.5)$$

Using the fact that $u_{\rho+\varepsilon\tilde{\rho}}$ and u_ρ are maximizers of (2.9) with electron density $\rho + \varepsilon\tilde{\rho}$ and ρ respectively, we have

$$\int_{\mathbb{R}^3} u_\rho(\rho + \varepsilon\tilde{\rho}) \leq \int_{\mathbb{R}^3} u_{\rho+\varepsilon\tilde{\rho}}(\rho + \varepsilon\tilde{\rho}) \quad \text{and} \quad -\int_{\mathbb{R}^3} u_\rho \rho \leq -\int_{\mathbb{R}^3} u_{\rho+\varepsilon\tilde{\rho}} \rho. \quad (3.6)$$

Substituting (3.6) into (3.5) gives

$$\int_{\mathbb{R}^3} u_\rho \tilde{\rho} \leq D_\varepsilon \leq \int_{\mathbb{R}^3} u_{\rho+\varepsilon\tilde{\rho}} \tilde{\rho}. \quad (3.7)$$

Under the uniqueness and continuity assumption (2.10), the right-hand side of (3.7) converges to the left-hand side as $\varepsilon \rightarrow 0$. Hence, for any $\tilde{\rho}$ with $\int_{\mathbb{R}^3} \tilde{\rho} = 0$, $\lim_{\varepsilon \rightarrow 0} D_\varepsilon$ exists and equals $\int_{\mathbb{R}^3} u_\rho \tilde{\rho}$. This together with definition (3.3) leads to $v_{\text{SCE}}[\rho] = u_\rho$.

Note that the map $\tilde{\rho} \mapsto \int_{\mathbb{R}^3} \frac{\delta V_{\text{ee}}^{\text{SCE}}[\rho]}{\delta \rho} \tilde{\rho}$ is unique up to an additive constant since $\int_{\mathbb{R}^3} \tilde{\rho} = 0$. Therefore, the functional derivative viewed as a function can be modified by any additive constant C . This completes the proof. \square

The Kantorovich potential u_ρ is related to the co-motion functions in the Monge formulation, as noted and justified in [14]. In what follows, we give a more mathematical derivation of this relation, which avoids the interpretation of the effective potential as a Lagrange multiplier and clarifies the relationship between the variational principle (3.13), introduced in [14], and the Kantorovich dual variational principle. Note that the interpretation of the effective potential as a Lagrange multiplier coming from the marginal constraint is heuristically correct, but difficult to make rigorous, the difficulties being related to the notorious “ v -representability-problem”, as will be discussed elsewhere.

Theorem 3.2. *Let ρ_N be the minimizer of the optimal transport problem (2.3) with given single-particle density ρ . If u_ρ is the Kantorovich potential, i.e., the maximizer of (2.9), and u_ρ is differentiable, then*

$$\nabla u_\rho(\mathbf{r}) = \nabla_{\mathbf{r}} c_{\text{ee}}(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) \quad \text{on } \text{supp}(\rho_N). \quad (3.8)$$

In particular, if ρ_N is of the Monge form (2.5), then

$$\nabla u_\rho(\mathbf{r}) = \nabla_{\mathbf{r}} c_{\text{ee}}(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) \big|_{\mathbf{r}_2=T_2(\mathbf{r}), \dots, \mathbf{r}_N=T_N(\mathbf{r})}. \quad (3.9)$$

Proof. We first note that $V_{\text{ee}}^{\text{SCE}}[\rho]$ is convex. To see this, let ρ be a convex combination $(1-t)\rho_A + t\rho_B$ for some $t \in (0,1)$, and ρ_N^A, ρ_N^B be the minimizers of (2.3) corresponding to the single-particle densities ρ_A and ρ_B , we have

$$V_{\text{ee}}^{\text{SCE}}[\rho] \leq (1-t)V_{\text{ee}}[\rho_N^A] + tV_{\text{ee}}[\rho_N^B] = (1-t)V_{\text{ee}}^{\text{SCE}}[\rho_A] + tV_{\text{ee}}^{\text{SCE}}[\rho_B]. \quad (3.10)$$

Since $V_{\text{ee}}^{\text{SCE}}[\rho]$ is convex, it equals its double Legendre transform. Denoting the Legendre transform of a functional F by F^* , we have

$$V_{\text{ee}}^{\text{SCE}*}[v] = \max_{\rho} \left(\int_{\mathbb{R}^3} v\rho - V_{\text{ee}}^{\text{SCE}}[\rho] \right).$$

By combining the maximization over ρ and minimization over $\rho_N \mapsto \rho$ in (2.3), we have

$$\begin{aligned} -V_{\text{ee}}^{\text{SCE}*}[v] &= \min_{\rho} \left(V_{\text{ee}}^{\text{SCE}}[\rho] - \int_{\mathbb{R}^3} v\rho \right) \\ &= \min_{\rho} \min_{\rho_N \mapsto \rho} \left(\int_{\mathbb{R}^{3N}} c_{\text{ee}} \rho_N - \int_{\mathbb{R}^3} v\rho \right) \\ &= \min_{\rho_N \mapsto \rho} \int_{\mathbb{R}^{3N}} \rho_N(\mathbf{r}_1, \dots, \mathbf{r}_N) \left(c_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) - \sum_{i=1}^N v(\mathbf{r}_i) \right). \end{aligned} \quad (3.11)$$

Then the double Legendre transform is

$$V_{\text{ee}}^{\text{SCE**}}[\rho] = \max_v \left(\int_{\mathbb{R}^3} v\rho - V_{\text{ee}}^{\text{SCE}*}[v] \right). \quad (3.12)$$

Combining (3.11), (3.12), and the fact that $V_{\text{ee}}^{\text{SCE}}$ equals its double Legendre transform results in the following variational principle

$$V_{\text{ee}}^{\text{SCE}}[\rho] = \max_v \left(\int_{\mathbb{R}^3} v\rho + \min_{\rho_N} \int_{\mathbb{R}^{3N}} (c_{\text{ee}} - \sum_{i=1}^N v(\mathbf{r}_i))\rho_N \right). \quad (3.13)$$

Note that the constraint $\rho_N \mapsto \rho$ has been eliminated in (3.13). For any fixed v , let $\mathcal{V}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{i=1}^N v(\mathbf{r}_i)$. The inner variational principle of (3.13) reads

$$\min_{\rho_N} \int_{\mathbb{R}^{3N}} (c_{\text{ee}} - \mathcal{V})\rho_N.$$

Since $c_{\text{ee}} - \mathcal{V}$ is a pure multiplicative operator, it follows (provided v is differentiable) that the support of any minimizer ρ_N must be contained in the set of absolute minimizers of $c_{\text{ee}} - \mathcal{V}$. Note that on the latter set, $\nabla(c_{\text{ee}} - \mathcal{V}) = 0$. Therefore, we have

$$\nabla_{\mathbf{r}_i}(c_{\text{ee}} - \mathcal{V}) = \nabla_{\mathbf{r}_i}c_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) - \nabla_{\mathbf{r}_i}v(\mathbf{r}_i) = 0 \quad \text{on } \text{supp}(\rho_N) \quad (3.14)$$

for $i = 1, \dots, N$.

According to Lemma A.1 in the appendix, if v_0 is a maximizer of (3.13), then the corresponding minimizer ρ_N^0 of the inner optimization of (3.13) is exactly the minimizer of the original problem (2.3). Therefore, (3.14) implies that if ρ_N^0 is a minimizer of (2.3), and v_0 is differentiable, then

$$\nabla_{\mathbf{r}_i}v_0(\mathbf{r}_i) = \nabla_{\mathbf{r}_i}c_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) \quad \text{on } \text{supp}(\rho_N^0), \quad i = 1, \dots, N. \quad (3.15)$$

In order to obtain (3.8), it is now only necessary to show that $u_\rho(\mathbf{r}) = v_0(\mathbf{r}) + \mu$ with some constant μ . Note that the maximum value of (3.13) is invariant under changing v by an additive constant, because the two integrals involving v cancel. Therefore, the maximization over v in (3.13) may be restricted to v 's with the additional property

$$\min_{(\mathbf{r}_1, \dots, \mathbf{r}_N) \in \mathbb{R}^{3N}} \left(c_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) - \sum_{i=1}^N v(\mathbf{r}_i) \right) = 0. \quad (3.16)$$

For these v 's, the minimization over ρ_N in (3.13) can be carried out explicitly (just place the support of ρ_N at the global minimizers of $c_{\text{ee}} - \mathcal{V}$). It then follows from (3.13) that

$$V_{\text{ee}}^{\text{SCE}}[\rho] = \max \left\{ \int_{\mathbb{R}^3} v\rho, \quad v \text{ satisfies (3.16)} \right\}. \quad (3.17)$$

Since $\int_{\mathbb{R}^3}(v+C)\rho$ is increasing as a function of the additive constant C , condition (3.16) can be changed into the inequality

$$\sum_{i=1}^N v(\mathbf{r}_i) \leq \sum_{1 \leq i < j \leq N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (3.18)$$

without affecting the maximal value in (3.17). This yields the Kantorovich dual form (2.9). Therefore, any maximizer of (3.13) satisfies that $u_\rho(\mathbf{r}) = v_0(\mathbf{r}) + \mu$ with some constant μ , which together with (3.15) implies (3.8).

If the minimizer of (2.3) is of the Monge form (2.5), then we have $\text{supp}(\rho_N) \subset M$ with M given by (2.6). This implies (3.9) and completes the proof. \square

Since the Coulomb cost function c_{ee} is given by (1.2), Eq. (3.9) is reduced to the following relation according to Theorem 3.2 (see also [14])

$$\nabla u_\rho(\mathbf{r}) = - \sum_{i=2}^N \frac{\mathbf{r} - T_i(\mathbf{r})}{|\mathbf{r} - T_i(\mathbf{r})|^3}. \quad (3.19)$$

In case $N = 2$, there is only one co-motion function T , and

$$\nabla u_\rho(\mathbf{r}) = - \frac{\mathbf{r} - T(\mathbf{r})}{|\mathbf{r} - T(\mathbf{r})|^3}. \quad (3.20)$$

Note that solving this equation for $T(\mathbf{r})$ gives an instance of the celebrated Gangbo-McCann formula [31] for the optimal map in terms of the Kantorovich potential. For the Coulomb cost this formula takes the form [21]

$$T(\mathbf{r}) = \mathbf{r} + \frac{\nabla u_\rho(\mathbf{r})}{|\nabla u_\rho(\mathbf{r})|^{3/2}}. \quad (3.21)$$

However, unlike (3.21), formula (3.20) generalizes (in the form of (3.19)) to many-body or multi-marginal problems. Thus formula (3.19) should be viewed as the correct generalization of the Gangbo-McCann formula to multi-marginal problems. We note that our derivation did not make use of Coulombic features of the cost; the same arguments yield a version of Eq. (3.19) for general pair costs of form $c_{ee}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{i < j} w(\mathbf{r}_i - \mathbf{r}_j)$, or Eq. (3.9) for fully general costs.

Using Theorem 3.1 and 3.2, we can derive the Kohn-Sham equations corresponding to the SCE energy functional (3.2) with a computable effective potential. It is the Euler-Lagrange equation corresponding to this minimization problem (after a unitary transformation to diagonalize the symmetric $N \times N$ matrix of Lagrange multipliers): find $\lambda_i \in \mathbb{R}$, $\phi_i \in H^1(\mathbb{R}^3)$ ($i = 1, 2, \dots, N$) such that

$$\begin{cases} \left(-\frac{1}{2}\Delta + v_{\text{ext}} + v_{\text{SCE}}[\rho_\Phi] \right) \phi_i &= \lambda_i \phi_i \quad \text{in } \mathbb{R}^3, \quad i = 1, 2, \dots, N, \\ \int_{\mathbb{R}^3} \phi_i \phi_j &= \delta_{ij}. \end{cases} \quad (3.22)$$

This is a nonlinear eigenvalue problem, where the potential $v_{\text{SCE}}[\rho_\Phi]$ depends on the electron density ρ_Φ associated with the orbitals ϕ_i . A self-consistent field (SCF) iteration algorithm is commonly resorted to for this nonlinear problem. In each iteration step of the algorithm, a new effective potential is constructed from a trial electron density and a linear eigenvalue problem is then solved to obtain the low-lying eigenvalues.

We shall comment further on the additive constant in (3.4). The above equations remain valid when $v_{\text{SCE}}[\rho_\Phi]$ is modified by an arbitrary additive constant. This yields the same Kohn-Sham orbitals ϕ_i and only leads to a corresponding shift of the nonlinear eigenvalues λ_i . However, as pointed out in [24], it is only when $v_{\text{SCE}}[\rho_\Phi]$ is precisely the Kantorovich potential that the ground state energy can equal the sum of Kohn-Sham eigenvalues, i.e.,

$$E_0 = \sum_{i=1}^N \lambda_i.$$

In summary, an SCF algorithm for solving the Kohn-Sham equation (3.22) is given by

Algorithm 3.3. *SCF iterations for SCE-based Kohn-Sham equations*

1. Given $\epsilon > 0$. Let $k = 0$ and ρ_0 be an initial electron density.
2. Calculate the co-motion functions from ρ_k (by using the numerical methods introduced in the next section).
3. Calculate the effective potential $v_{\text{SCE}}[\rho_k]$ by (3.19).
4. Solve the linear eigenvalue problem

$$\left(-\frac{1}{2}\Delta + v_{\text{ext}} + v_{\text{SCE}}[\rho_k]\right)\phi_i = \lambda_i\phi_i \quad i = 1, \dots, N$$

for low-lying eigenvalues to obtain new Kohn-Sham orbitals, from which a new electron density ρ_k^{out} can be calculated.

5. If $\|\rho_k - \rho_k^{\text{out}}\| < \epsilon$, stop; else, generate a new electron density ρ_{k+1} by some charge mixing technique and go to 2.

4 Numerical discretizations of optimal transportation

In each iteration of the SCF algorithm for solving (3.22), one has to construct $v_{\text{SCE}}[\tilde{\rho}]$ from a trial electron density $\tilde{\rho}$. According to (3.19), this requires the solution of the optimal transport problem (2.3) with a given single-particle density to obtain the co-motion functions T_i , $i = 2, \dots, N$. For simplicity, we only consider the case $N = 2$, where only one co-motion function has to be calculated (which is denoted by T in the following). For systems with more than two electrons, we refer to Section 6 for a future perspective.

We discretize the computational domain into n finite elements e_1, \dots, e_n . (We replace \mathbb{R}^3 by a bounded domain so that it can be discretized into a finite number of elements. This is reasonable since the electron density $\rho(\mathbf{r})$ of a confined system decays exponentially fast to zero as $|\mathbf{r}| \rightarrow \infty$ [32].) Each element is represented by a point \mathbf{a}_k located at its barycenter and its electron mass $\rho_k = \int_{e_k} \rho(\mathbf{r}) d\mathbf{r}$. Within this discretization, we can approximate the two-particle density $|\Psi(\mathbf{r}_1, \mathbf{r}_2)|^2$ by a matrix $X = (x_{kl}) \in \mathbb{R}^{n \times n}$ with $x_{kl} = |\Psi(\mathbf{a}_k, \mathbf{a}_l)|^2$. (Alternatively, one could identify the entries with the average $x_{kl} = \frac{1}{|e_k| \cdot |e_l|} \int_{e_k} \int_{e_l} |\Psi(\mathbf{r}_1, \mathbf{r}_2)|^2 d\mathbf{r}_1 d\mathbf{r}_2$.) The continuous problem (2.3) is then discretized into

$$\begin{aligned} \min_X \quad & \sum_{1 \leq k, l \leq n} \frac{x_{kl}}{|\mathbf{a}_k - \mathbf{a}_l|} \\ \text{s.t.} \quad & \sum_{1 \leq k \leq n} x_{kl} = \frac{1}{2}\rho_k, \quad l = 1, \dots, n \\ & \sum_{1 \leq l \leq n} x_{kl} = \frac{1}{2}\rho_l, \quad k = 1, \dots, n \\ & x_{kl} \geq 0. \end{aligned} \tag{4.1}$$

Note that (4.1) is a linear programming problem of the form

$$\begin{aligned} \min_x & f^T x \\ \text{s.t.} & Ax = b \text{ and } x_k \geq 0, \end{aligned}$$

where x is the vector containing the entries of X . We can solve this problem by standard optimization routines like ‘*linprog*’ in *Matlab*. Due to the symmetry of the problem, one can assume that $x_{lk} = x_{kl}$ and only needs to consider x_{kl} for $k \leq l$.

As a remark, the dual problem of (4.1) (in the sense of linear programming) results in a discretized version of the Kantorovich dual formulation (2.9).

The solution of (4.1) entails an approximation of the co-motion functions at the barycenters $\{\mathbf{a}_k\}_{1 \leq k \leq n}$ via the matrix $X = (x_{kl})$:

$$T_n(\mathbf{a}_k) = \sum_{l=1}^n \mathbf{a}_l \frac{2x_{kl}}{\rho_l}, \quad k = 1, \dots, n, \quad (4.2)$$

where x_{kl} can also be regarded as the mass of electron transported from \mathbf{a}_k to \mathbf{a}_l . If the discretization is sufficiently fine, i.e., n large enough, then T_n is a good approximation of T (see the following numerical example).

For a *uniform* discretization $\{e_k\}_{1 \leq k \leq n}$, the degrees of freedom for linear programming (4.1) may be huge. To reduce the computational cost, we use a locally refined mesh instead, which has more elements where the electron density is high and less elements where the electron density is low. Generally speaking, the optimal mesh may be such that each element e_k has almost equal electron mass ρ_k . This type of mesh can be generated by an adaptive procedure, say, one refines the element when its electron mass is larger than a given threshold and coarse it otherwise. Since the electron density decays exponentially fast to zero as $|\mathbf{r}| \rightarrow \infty$, the mesh is much coarser far away from the nuclei than close to the nuclei, which reduces the degrees of freedom significantly.

As a remark, let us assume for a moment that all elements have exactly the same mass, $\rho_k = \bar{\rho}$ for all k . Then the constraints in (4.1) force X to be a *doubly stochastic* matrix (up to a global scaling factor) with nonnegative entries. According to Birkhoff’s theorem, the extremal points of the convex set of admissible matrices X are the permutations, i.e., matrices with exactly one non-zero entry $\frac{1}{2}\bar{\rho}$ in each row (or column). Since the optimum is obtained at an extremal point, the optimizer can be chosen of this form. We have thus derived a discrete analogue of the Monge formulation, since the sum on the right of Eq. (4.2) will have exactly one nonzero term.

Another important technique to reduce the computational cost is to exploit the symmetry of the system. If the electron density ρ has some kind of symmetric property, then we can reduce the computations to some subdomain accordingly. For example, [14] gives an explicit formula of co-motion functions for spherically symmetric electron densities by making use of the symmetry. More precisely, it is proven that if the density has the form $\rho(\mathbf{r}) = h(|\mathbf{r}|)$ with some function $h : [0, \infty) \rightarrow \mathbb{R}$, then the corresponding co-motion function T has to be spherically symmetric itself, that is

$$T(\mathbf{r}) = g(|\mathbf{r}|) \frac{\mathbf{r}}{|\mathbf{r}|}, \quad \forall \mathbf{r} \in \mathbb{R}^3$$

with some function $g : [0, \infty) \rightarrow \mathbb{R}$. This reduces the three-dimensional spherically symmetric problem into a one-dimensional problem.

Here we consider cylindrically symmetric systems, for instances, biatomic molecules. The following theorem states that the co-motion function T inherits the cylindrical symmetry of the density.

Theorem 4.1. *Let $N = 2$ and denote the cylindrical coordinates by (γ, φ, z) . If $\rho(\mathbf{r}) = \varrho(\gamma, z)$ with some function $\varrho : [0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}$, then the corresponding co-motion function T satisfies*

$$T : (\gamma, \varphi, z) \mapsto (\gamma', \varphi + \pi, z') \quad \forall (\gamma, z) \in [0, \infty) \times \mathbb{R}, \quad (4.3)$$

where $(\gamma', z') = \ell(\gamma, z)$ with some map $\ell : [0, \infty) \times \mathbb{R} \rightarrow [0, \infty) \times \mathbb{R}$.

Proof. Let \mathcal{R}_θ be the rotation operator with angle θ around the z-axis, i.e., $\mathcal{R}_\theta(\gamma, \varphi, z) = (\gamma, \varphi + \theta, z)$. Let ρ_2 be a minimizer of (2.3) with single-particle electron density ρ , such that $\rho_2(\mathbf{r}_1, \mathbf{r}_2) = \frac{\rho(\mathbf{r}_1)}{2} \delta(\mathbf{r}_2 - T(\mathbf{r}_1))$ with T the corresponding co-motion function.

Let $\tilde{\rho}_2(\mathbf{r}_1, \mathbf{r}_2) = \rho_2(\mathcal{R}_\theta \mathbf{r}_1, \mathcal{R}_\theta \mathbf{r}_2)$. We claim that $\tilde{\rho}_2$ is also a minimizer of (2.3). To see this, we observe that

$$\tilde{\rho}_2(\mathbf{r}_1, \mathbf{r}_2) = \frac{\rho(\mathcal{R}_\theta \mathbf{r}_1)}{2} \delta(\mathcal{R}_\theta \mathbf{r}_2 - T(\mathcal{R}_\theta \mathbf{r}_1)),$$

which satisfies the marginal constraint

$$\tilde{\rho}_2 \mapsto \rho \quad (4.4)$$

since $\rho(\mathbf{r}) = \rho(\mathcal{R}_\theta \mathbf{r})$. Moreover, we have

$$\int_{\mathbb{R}^6} \frac{\tilde{\rho}_2(\mathbf{r}_1, \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 = \int_{\mathbb{R}^6} \frac{\rho_2(\mathbf{r}_1, \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \quad (4.5)$$

from the fact that the cost function $\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}$ is invariant under the map $(\mathbf{r}_1, \mathbf{r}_2) \rightarrow (\mathcal{R}_\theta \mathbf{r}_1, \mathcal{R}_\theta \mathbf{r}_2)$. (4.4) and (4.5) together imply that $\tilde{\rho}_2$ is a minimizer of (2.3).

Therefore, $\tilde{T} : \mathbf{r} \mapsto \mathcal{R}_\theta^{-1} T(\mathcal{R}_\theta \mathbf{r})$ is also a co-motion function of this problem. Because the co-motion function is unique in the case of two particles (see [21]), we have

$$T(\mathcal{R}_\theta \mathbf{r}) = \mathcal{R}_\theta T(\mathbf{r}). \quad (4.6)$$

Since we minimize

$$\begin{aligned} \int_{\mathbb{R}^3} \frac{\rho(\mathbf{r})}{|\mathbf{r} - T(\mathbf{r})|} d\mathbf{r} &= \int_0^\infty \gamma d\gamma \int_{-\infty}^\infty dz \int_0^{2\pi} d\varphi \frac{\rho((\gamma, \varphi, z))}{|(\gamma, \varphi, z) - T((\gamma, \varphi, z))|} d\varphi \\ &\stackrel{\text{fix } \varphi}{=} \int_0^\infty \gamma d\gamma \int_{-\infty}^\infty dz \int_0^{2\pi} d\theta \frac{\rho(\mathcal{R}_\theta(\gamma, \varphi, z))}{|\mathcal{R}_\theta(\gamma, \varphi, z) - T(\mathcal{R}_\theta(\gamma, \varphi, z))|} d\theta \\ &\stackrel{(4.6)}{=} \int_0^\infty \gamma d\gamma \int_{-\infty}^\infty dz \int_0^{2\pi} d\theta \frac{\rho(\mathcal{R}_\theta(\gamma, \varphi, z))}{|\mathcal{R}_\theta(\gamma, \varphi, z) - \mathcal{R}_\theta T((\gamma, \varphi, z))|} d\theta \\ &= 2\pi \int_0^\infty \int_{-\infty}^\infty \frac{\gamma \varrho((\gamma, z))}{|(\gamma, \varphi, z) - T((\gamma, \varphi, z))|} d\gamma dz, \end{aligned}$$

T has to satisfy (4.3) to maximize the denominator. This completes the proof. \square

Example. We take a one-dimensional two-electron system as an example to illustrate our numerical method. Although this type of algorithm is not particularly interesting for one-dimensional systems since analytical formulations are known [27, 24], it is more suitable to present the numerical results and explain our idea.

Let $\Omega = [-5, 5]$ and $\rho(x) = 0.4 - 0.08|x|$, see Figure 4.1a. The exact co-motion function can be calculated explicitly

$$T(x) = \begin{cases} 5 \left(1 - [1 - 0.5(x+5)(0.4 + 0.08x)]^{1/2} \right) & \text{if } x \leq 0, \\ -5 \left(1 - [1 - 0.5(-x+5)(0.4 - 0.08x)]^{1/2} \right) & \text{if } x > 0. \end{cases} \quad (4.7)$$

We observe in Figure 4.1b that the numerical approximations of the co-motion function can be very accurate (compared with the exact formula (4.7)). Figure 4.1c shows the convergence of the uniform numerical approximations. The nonuniform mesh (red elements in Figure 4.1a) achieves a much higher accuracy with the same degrees of freedom, see Figure 4.1d. Actually, we observe that the errors of a nonuniform mesh with $n = 20$ is even smaller compared to a uniform mesh with $n = 40$ on average.

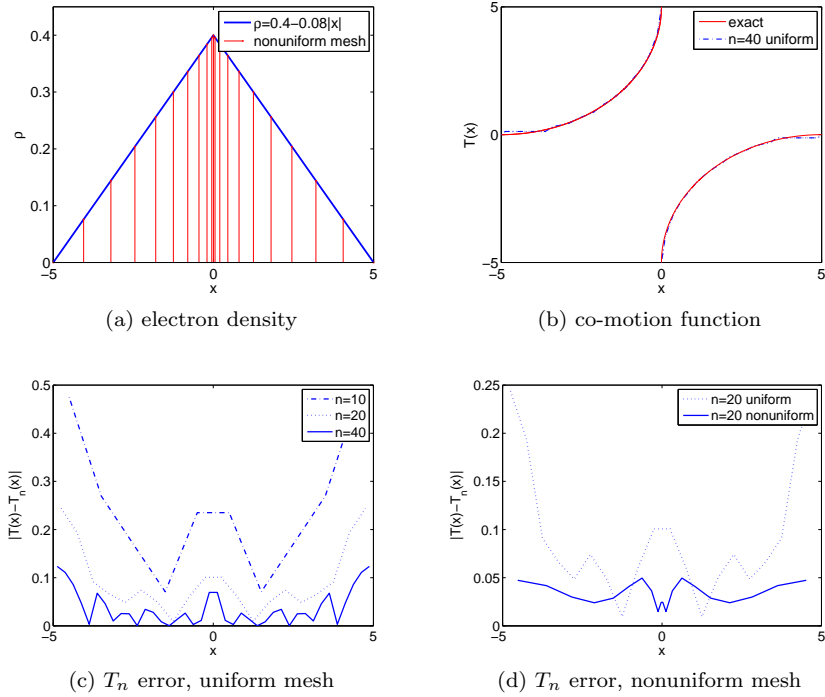


Figure 4.1: (a) The one-dimensional electron density ρ and a nonuniform discretization. (b) The co-motion function corresponding to ρ and its approximation. (c) Numerical errors of T_n using uniform meshes, and (d) a nonuniform mesh.

Remember that the aforementioned electron density is symmetric in the

sense of

$$\rho(x) = \rho(-x).$$

Therefore, the corresponding co-motion function is symmetric itself, i.e.

$$T(x) = -T(-x),$$

which maps $[-5, 0]$ to $[0, 5]$ and maps $[0, 5]$ to $[-5, 0]$. Hence, it is only necessary to calculate $T(x)$ on half of the domain, say, $[-5, 0]$.

An important application of the numerical methods introduced above is to simulate the H_2 molecule at its dissociating limit. We provide more details in the next section.

5 H_2 bond disassociation

We consider H_2 molecule in this section. Let $R > 0$ and $\mathbf{R}_A = (-R, 0, 0)$, $\mathbf{R}_B = (R, 0, 0)$ be the locations of two hydrogen atoms. Physically, the hydrogen molecule should dissociate into two free hydrogen atoms as the bond length $2R \rightarrow \infty$, with the ground state spin-unpolarized. The spin-restricted Hartree-Fock and Kohn-Sham DFT models give the correct spin multiplicity, but overestimate total energies, i.e., higher than that of two free hydrogen atoms. In comparison, the spin-unrestricted models give fairly good total energies, while the wave functions are spin-contaminated, which is known as “symmetry breaking” in H_2 bond dissociation.

Here we focus on the SCE-DFT model without symmetry breaking, and show both theoretically and numerically that the restricted Kohn-Sham model (3.1) gives the correct ground state energy in the dissociation limit $R \rightarrow \infty$. Denote by e_0 the ground state energy of a single hydrogen atom

$$e_0 = \inf \left\{ \frac{1}{2} \int_{\mathbb{R}^3} |\nabla \phi(\mathbf{r})|^2 d\mathbf{r} - \int_{\mathbb{R}^3} \frac{|\phi(\mathbf{r})|^2}{|\mathbf{r}|} d\mathbf{r}, \phi \in H^1(\mathbb{R}^3), \|\phi\|_{L^2(\mathbb{R}^3)} = 1 \right\}. \quad (5.1)$$

$E_{\text{SCE}}(R)$ denotes the ground state energy of the hydrogen molecule in the SCE-DFT model (3.2)

$$E_{\text{SCE}}(R) = \frac{1}{2R} + \inf \left\{ \int_{\mathbb{R}^3} |\nabla \phi(\mathbf{r})|^2 d\mathbf{r} + 2 \int_{\mathbb{R}^3} v_{\text{ext}}(\mathbf{r}) |\phi(\mathbf{r})|^2 d\mathbf{r} + V_{\text{ee}}^{\text{SCE}}[2|\phi|^2], \right. \\ \left. \phi \in H^1(\mathbb{R}^3), \|\phi\|_{L^2(\mathbb{R}^3)} = 1 \right\}, \quad (5.2)$$

where $v_{\text{ext}}(\mathbf{r}) = -\frac{1}{|\mathbf{r} - \mathbf{R}_A|} - \frac{1}{|\mathbf{r} - \mathbf{R}_B|}$. The following result indicates that the SCE-DFT model is correct for the H_2 molecule at its dissociating limit.

Theorem 5.1. *Let e_0 and $E_{\text{SCE}}(R)$ be given by (5.1) and (5.2) respectively. We have*

$$\lim_{R \rightarrow \infty} E_{\text{SCE}}(R) = 2e_0. \quad (5.3)$$

Proof. First, we establish an upper bound of $E_{\text{SCE}}(R)$. Let $\varphi(r) = e^{-r}/\sqrt{\pi}$ and

$$\psi(\mathbf{r}) = \left(\frac{1}{2} \left(\varphi^2(|\mathbf{r} - \mathbf{R}_A|) + \varphi^2(|\mathbf{r} - \mathbf{R}_B|) \right) \right)^{1/2}.$$

Note that φ is the minimizer of (5.1), and $\|\varphi\|_{L^2(\mathbb{R}^3)} = 1$ implies $\|\psi\|_{L^2(\mathbb{R}^3)} = 1$. We have

$$E_{\text{SCE}}(R) \leq \frac{1}{2R} + \int_{\mathbb{R}^3} |\nabla \psi(\mathbf{r})|^2 d\mathbf{r} + 2 \int_{\mathbb{R}^3} v_{\text{ext}}(\mathbf{r}) |\psi(\mathbf{r})|^2 d\mathbf{r} + V_{\text{ee}}^{\text{SCE}}[2|\psi|^2]. \quad (5.4)$$

Let $\phi_1(\mathbf{r}) = \varphi(|\mathbf{r} - \mathbf{R}_A|)$ and $\phi_2(\mathbf{r}) = \varphi(|\mathbf{r} - \mathbf{R}_B|)$. A direct calculation leads to

$$\begin{aligned} \int_{\mathbb{R}^3} |\nabla \psi|^2 &= \int_{\mathbb{R}^3} \frac{|\phi_1 \nabla \phi_1 + \phi_2 \nabla \phi_2|^2}{4(\phi_1^2 + \phi_2^2)} \\ &\leq \frac{1}{2} \int_{\mathbb{R}^3} (|\nabla \phi_1|^2 + |\nabla \phi_2|^2) = \int_{\mathbb{R}^3} |\nabla \varphi|^2 \end{aligned} \quad (5.5)$$

and

$$\begin{aligned} 2 \int_{\mathbb{R}^3} v_{\text{ext}}(\mathbf{r}) |\psi(\mathbf{r})|^2 d\mathbf{r} &= - \int_{\mathbb{R}^3} \frac{\phi_1^2(\mathbf{r}) + \phi_2^2(\mathbf{r})}{|\mathbf{r} - \mathbf{R}_A|} d\mathbf{r} - \int_{\mathbb{R}^3} \frac{\phi_1^2(\mathbf{r}) + \phi_2^2(\mathbf{r})}{|\mathbf{r} - \mathbf{R}_B|} d\mathbf{r} \\ &= -2 \int_{\mathbb{R}^3} \frac{\varphi^2(|\mathbf{r}|)}{|\mathbf{r}|} d\mathbf{r} - \int_{\mathbb{R}^3} \frac{\phi_2^2(\mathbf{r})}{|\mathbf{r} - \mathbf{R}_A|} d\mathbf{r} - \int_{\mathbb{R}^3} \frac{\phi_1^2(\mathbf{r})}{|\mathbf{r} - \mathbf{R}_B|} d\mathbf{r} \\ &\leq -2 \int_{\mathbb{R}^3} \frac{\varphi^2(|\mathbf{r}|)}{|\mathbf{r}|} d\mathbf{r}. \end{aligned} \quad (5.6)$$

Let $\rho_2(\mathbf{r}_1, \mathbf{r}_2) = 2|\psi(\mathbf{r}_1)|^2 \delta(\mathbf{r}_2, \mathbf{r}_1 - \mathbf{R}_A + \mathbf{R}_B)$, we have

$$V_{\text{ee}}^{\text{SCE}}[2|\psi|^2] \leq \int_{\mathbb{R}^6} \frac{\rho_2(\mathbf{r}_1, \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 = \int_{\mathbb{R}^3} \frac{2|\psi(\mathbf{r}_1)|^2}{|\mathbf{R}_A - \mathbf{R}_B|} d\mathbf{r}_1 = \frac{1}{R}. \quad (5.7)$$

Taking (5.4), (5.5), (5.6) and (5.7) into account, we have

$$E_{\text{SCE}}(R) \leq 2e_0 + \frac{3}{2R}. \quad (5.8)$$

To give a lower bound of $E_{\text{SCE}}(R)$, we observe that

$$\begin{aligned} E_{\text{SCE}}(R) &\geq \frac{1}{2R} + \inf_{\|\phi\|_{L^2}=1} \left\{ \int_{\mathbb{R}^3} |\nabla \phi(\mathbf{r})|^2 d\mathbf{r} + 2 \int_{\mathbb{R}^3} v_{\text{ext}}(\mathbf{r}) |\phi(\mathbf{r})|^2 d\mathbf{r} \right\} \\ &= \frac{1}{2R} + 2 \inf_{\|\phi\|_{L^2}=1} \langle \phi | \hat{h} | \phi \rangle, \end{aligned} \quad (5.9)$$

where \hat{h} is the H_2^+ Hamiltonian $\hat{h} = -\frac{1}{2}\Delta + v_{\text{ext}}$. We claim that

$$\langle \phi | \hat{h} | \phi \rangle \geq (e_0 - O(R^{-1})) \|\phi\|_{L^2}^2 \quad \forall \phi \in H^1(\mathbb{R}^3). \quad (5.10)$$

To show (5.10), we first decompose the unity function on \mathbb{R} into two smooth cutoff functions $\tilde{\zeta}_1$ and $\tilde{\zeta}_2$, such that $\tilde{\zeta}_1^2 + \tilde{\zeta}_2^2 = 1$, $\tilde{\zeta}_1(x) = 0$ for $x < -\frac{1}{2}$, $\tilde{\zeta}_2(x) = 0$ for $x > \frac{1}{2}$, and $|\nabla \tilde{\zeta}_i| \leq C^*$ with some constant C^* . Let

$$\zeta_i(\mathbf{r}) = \zeta_i(x, y, z) = \tilde{\zeta}_i(x/R) \quad i = 1, 2$$

and $\phi_i = \zeta_i \phi$. We have $|\nabla \zeta_i| \leq C^*/R$ and

$$\sum_{i=1}^2 |\nabla \phi_i|^2 = \sum_{i=1}^2 (|\nabla \zeta_i|^2) |\phi|^2 + |\nabla \phi|^2.$$

Therefore,

$$\begin{aligned}\langle \phi | \hat{h} | \phi \rangle &= \frac{1}{2} \sum_{i=1}^2 \int_{\mathbb{R}^3} |\nabla \phi_i|^2 - \int_{\mathbb{R}^3} \sum_{i=1}^2 (|\nabla \zeta_i|^2) |\phi|^2 + \int_{\mathbb{R}^3} v_{\text{ext}} (|\phi_1|^2 + |\phi_2|^2) \\ &\geq \sum_{i=1}^2 \int_{\mathbb{R}^3} \left(\frac{1}{2} |\nabla \phi_i|^2 - \frac{|\phi_i|^2}{|\mathbf{r}|} \right) - \left(\frac{2C^*}{R} \right)^2 - \frac{2}{R},\end{aligned}$$

which implies (5.10). Therefore, we obtain from (5.9) and (5.10) that

$$E_{\text{SCE}}(R) \geq 2e_0 + O(R^{-1}). \quad (5.11)$$

Together with the upper bound (5.8), this leads to (5.3) at the limit $R \rightarrow \infty$. \square

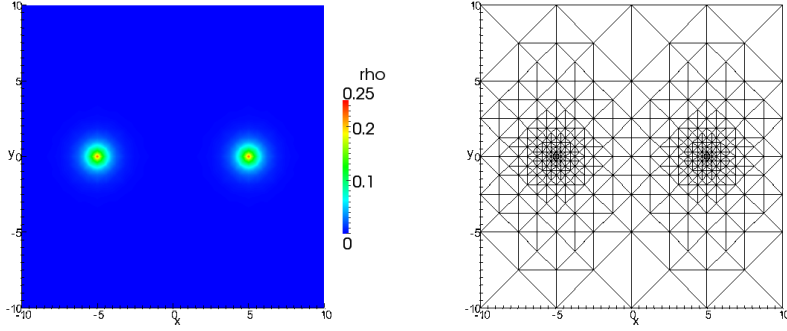


Figure 5.1: The electron density and corresponding mesh on slice $z = 0$ ($R = 5$).

In what follows, we present a numerical simulation of the dissociating H_2 molecule to support the theory. The computations are carried out on a bounded domain $\Omega = [-10, 10]^3$. We use Algorithm 3.3 to solve (3.22) for the ground state energies and electron densities for different bond length $2R$. Concerning the optimal transport problem, we use the numerical methods introduced in Section 4 and calculate the co-motion function in each SCF iteration step. The nonuniform mesh is generated by the package PHG [33], a toolbox for parallel adaptive finite element programs developed at the State Key Laboratory of Scientific and Engineering Computing of the Chinese Academy of Sciences. While the problem is effectively two-dimensional according to Theorem 4.1, we have performed the calculations in three dimensions since PHG is tailored to three-dimensional problems. Figure 5.1 shows a contour plot of the electron density at slice $z = 0$ and the corresponding mesh. One observes that the grid reflects the higher density around the nuclei.

To further reduce the computational cost, we can exploit the cylindrical symmetry of the system with the help of Theorem 4.1. As shown in Figure 5.2, the degrees of freedom can be reduced to 1/16 of the original volume. For the linear programming problem (4.1), we resort to MOSEK [34], a high-performance software for large-scale optimization problems.

The computational results are presented in Figure 5.3, in which we compare the bond energies in dependence of R using the LDA and SCE Kohn-Sham

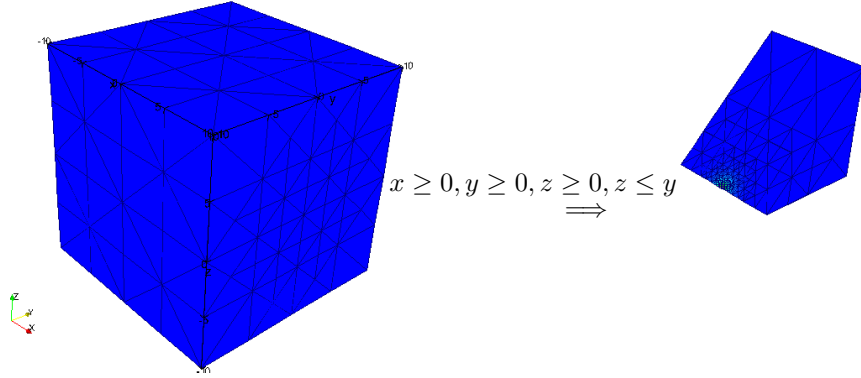


Figure 5.2: Symmetric decomposition of the computation domain Ω for H_2 .

methods, respectively. Here, the bond energies are the ground state energies of the systems minus $2e_0$, which is expected to be zero when the two hydrogen atoms are disassociated. Note that the SCE model shows the correct asymptotic behavior, while the LDA model fails at large R by giving too large energies. For comparison, the LDA error 0.065 a.u. in [35] for the infinitely stretched H_2 molecule is lower than in Figure 5.3 since twice the LDA-hydrogen energy is subtracted instead of twice the exact e_0 , but remains significant; errors of similar magnitude are reported there for other functionals such as B3LYP or PBE.

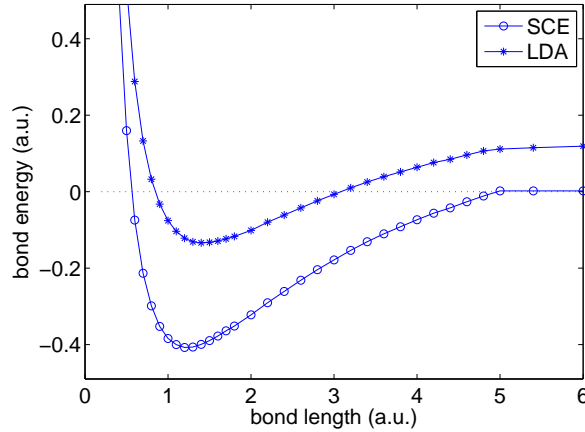


Figure 5.3: H_2 potential energy curve as a function of the bond length for the SCE and LDA models.

As physical explanations for these results, at long internuclear separations, if one electron is located near atom A, the other will be found close to atom B. This correlation is correctly reflected by the optimal transport model, hence the SCE model gives asymptotically the product of hydrogen orbitals on the two nuclei. In contrast, within the Kohn-Sham LDA framework, the two electrons

are constrained to be in the same spatial orbital and each electron experiences only the average effect of the other, thus each electron has equal probability of being near A or B, irrespective of the position of the other electron. The possibility of both electrons being on the same atom is not excluded, as reflected in the wrong asymptotic behavior of the disassociation energy in Figure 5.3.

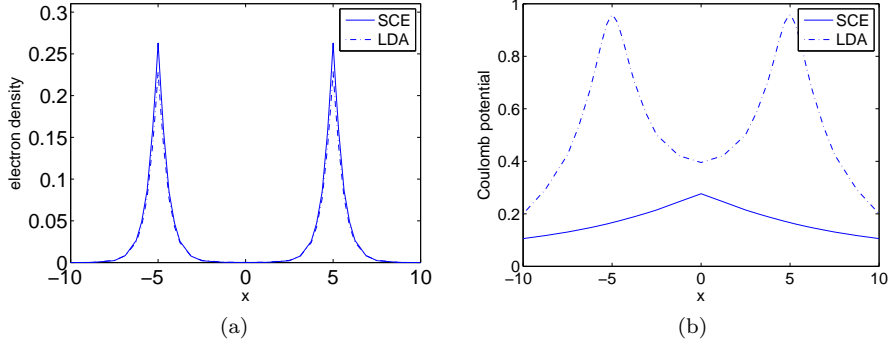


Figure 5.4: The SCE and LDA electron densities (a) and potentials (b) of H₂, plotted along the molecular axis.

We further compare the electron densities and Coulomb potentials obtained by the two different models in Figure 5.4. Note that the scalar offset of the potential in the figure is determined by the boundary conditions, i.e., $\frac{N}{R}$ for LDA and $\frac{N-1}{R}$ for SCE [14]. The electron densities are actually quite close, while the shape of the potentials differs substantially when R is large. The SCE potential is larger between the hydrogen atoms, favoring a depletion of the bond charge whenever the two atoms separate. This produces the correct results for large R .

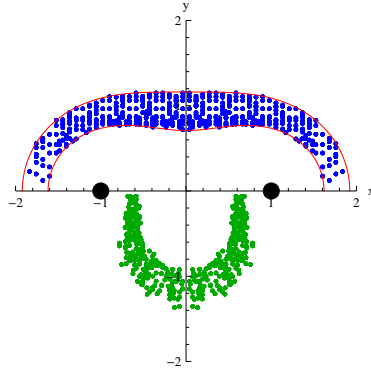


Figure 5.5: Optimal transport mapping of the region $0.04 \leq \rho(\mathbf{r}) \leq 0.08$ (indicated by the blue dots inside the red contours) to the green area, for the H₂ molecule with $R = 1$. Each blue dot corresponds to one barycenter in the numerical discretization, and has been rotated into the x - y -plane with $y \geq 0$ for visual clarity. The green dots are precisely the images of the blue dots under the optimal transport map.

Finally, as illustration of the co-motion function or optimal map, Figure 5.5 shows the image of the map on a density contour.

6 Conclusions and perspectives

The numerical discretization of the SCE optimal transport problem with Coulomb cost and two marginals leads to the linear programming problem (4.1), which can indeed be solved in practice, as we have demonstrated in a proof of concept calculation. The self-consistent SCE-DFT simulation of the dissociating H_2 molecule agrees well with the physically correct limit, unlike standard DFT models like LDA.

The theory of this paper applies to more general systems with arbitrary numbers of electrons, however, the numerical algorithms need further developments. Specifically, we can restrict the N -particle density to the ansatz

$$\rho_N(\mathbf{r}_1, \dots, \mathbf{r}_N) = \frac{\rho(\mathbf{r}_1)}{N} \gamma_2(\mathbf{r}_1, \mathbf{r}_2) \gamma_3(\mathbf{r}_1, \mathbf{r}_3) \cdots \gamma_N(\mathbf{r}_1, \mathbf{r}_N),$$

where ρ is the given single-particle density. Here, $\gamma_j(\mathbf{r}_1, \mathbf{r}_j)$ represents the probability of the j th electron being found at \mathbf{r}_j while the first electron is located at \mathbf{r}_1 . We have

$$\int_{\mathbb{R}^3} \gamma_j(\mathbf{r}_1, \mathbf{r}_j) d\mathbf{r}_j = 1 \quad j = 2, \dots, N.$$

With a given discretization $\{e_k\}_{1 \leq k \leq n}$ and barycenters \mathbf{a}_k of e_k , we can approximate $\gamma_j(\mathbf{r}_1, \mathbf{r}_j)$ by a matrix $X_j = (x_{j,kl}) \in \mathbb{R}^{n \times n}$ for $j = 2, \dots, N$. Then the continuous model (2.3) is reduced to

$$\begin{aligned} \min_{X_2, \dots, X_N} \quad & \sum_{1 \leq j \leq N} \sum_{k,l=1}^n \frac{x_{j,kl}}{|\mathbf{a}_k - \mathbf{a}_l|} \cdot \frac{\rho_k}{N} + \sum_{1 \leq i < j \leq N} \sum_{k,l,l'=1}^n \frac{x_{i,kl} \cdot x_{j,kl'}}{|\mathbf{a}_l - \mathbf{a}_{l'}|} \cdot \frac{\rho_k}{N} \\ \text{s.t.} \quad & \sum_{k=1}^n x_{i,kl} = 1, \quad l = 1, \dots, n, \quad i = 2, \dots, N \\ & \sum_{l=1}^n x_{i,kl} = 1, \quad k = 1, \dots, n, \quad i = 2, \dots, N \\ & x_{i,kl} \geq 0. \end{aligned} \tag{6.1}$$

This is a quadratic programming problem of the form

$$\begin{aligned} \min_x \quad & x^T H x + f^T x \\ \text{s.t.} \quad & A x = b \quad \text{and} \quad x_k \geq 0. \end{aligned}$$

By solving the above quadratic programming problem, we approximate the co-motion functions T_2, \dots, T_N by the matrices $X_2 = (x_{2,kl}), \dots, X_N = (x_{N,kl})$. Similar to (4.2), the co-motion functions can be approximated by

$$T_i(\mathbf{a}_k) \approx \sum_{1 \leq l \leq n} \mathbf{a}_l x_{i,kl} \quad k = 1, \dots, n, \quad i = 2, \dots, N.$$

However, a serious difficulty in solving (6.1) stems from the non-convexity of the matrix H . Moreover, the symmetric decomposition is not clear for systems with more than two electrons. We plan to investigate these issues in future work.

Acknowledgments C.M. acknowledges support from the DFG project FR 1275/3-1.

A Appendix

Lemma A.1. *If v_0 is a maximizer of (3.13) with single-particle density ρ , and ρ_N^0 is a minimizer of the inner optimization of (3.13), i.e. the minimizer of*

$$\min_{\rho_N} \int_{\mathbb{R}^{3N}} \left(c_{ee}(\mathbf{r}_1, \dots, \mathbf{r}_N) - \sum_{i=1}^N v_0(\mathbf{r}_i) \right) \rho_N(\mathbf{r}_1, \dots, \mathbf{r}_N) d\mathbf{r}_1 \dots d\mathbf{r}_N, \quad (\text{A.1})$$

then ρ_N^0 is exactly the minimizer of the constraint minimization problem (2.3).

Proof. Since v_0 is a maximizer of (3.13), we have

$$0 = \frac{d}{d\varepsilon} \bigg|_{\varepsilon=0} \left(\int_{\mathbb{R}^3} (v_0 + \varepsilon \tilde{v}) \rho + \min_{\rho_N} \int_{\mathbb{R}^{3N}} (c_{ee} - V_0 - \varepsilon \tilde{V}) \rho_N \right) \quad (\text{A.2})$$

for any \tilde{v} , where $V_0(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{i=1}^N v_0(\mathbf{r}_i)$ and $\tilde{V}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{i=1}^N \tilde{v}(\mathbf{r}_i)$.

We view the minimization over ρ_N in (A.2) as a quadratic variational problem for (spinless bosonic) normalized wavefunctions Φ (by identifying $|\Phi|^2 = \rho_N$):

$$\min_{\rho_N} \int_{\mathbb{R}^{3N}} (c_{ee} - V_0 - \varepsilon \tilde{V}) \rho_N = \min_{\Phi} \langle \Phi | c_{ee} - V_0 - \varepsilon \tilde{V} | \Phi \rangle.$$

Using first order perturbation theory and the fact that ρ_N^0 is the minimizer of $\min_{\rho_N} \int_{\mathbb{R}^{3N}} (c_{ee} - V_0) \rho_N$, we obtain

$$\min_{\rho_N} \int_{\mathbb{R}^{3N}} (c_{ee} - V_0 - \varepsilon \tilde{V}) \rho_N = \int_{\mathbb{R}^{3N}} (c_{ee} - V_0) \rho_N^0 + \varepsilon \int_{\mathbb{R}^{3N}} \tilde{V} \rho_N^0 + O(\varepsilon^2). \quad (\text{A.3})$$

Substituting (A.3) into (A.2) yields

$$\begin{aligned} 0 &= \int_{\mathbb{R}^{3N}} \tilde{V} \rho_N^0 - \int_{\mathbb{R}^3} \tilde{v} \rho \\ &= - \int_{\mathbb{R}^3} \tilde{v}(\mathbf{r}) \left(\rho(\mathbf{r}) - N \int_{\mathbb{R}^{3(N-1)}} \rho_N^0(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) d\mathbf{r}_2 \dots d\mathbf{r}_N \right) d\mathbf{r}, \end{aligned}$$

which indicates $\rho_N^0 \mapsto \rho$. That is, if v_0 is a maximizer of the outer optimization of (3.13), then the minimizer of the associated inner optimization in (3.13) automatically has the single-particle density ρ .

Moreover, since the term $\int_{\mathbb{R}^{3N}} V_0 \rho_N$ only depends on the single-particle density of ρ_N , ρ_N^0 must minimize $\int_{\mathbb{R}^{3N}} c_{ee} \rho_N$ under the constraint $\rho_N \mapsto \rho$ (because any other minimizer of (2.3) gives the same value for $\int_{\mathbb{R}^{3N}} (c_{ee} - V) \rho_N$ as ρ_N^0). Therefore, ρ_N^0 is also a minimizer of (2.3), which completes the proof. \square

References

- [1] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev. B*, 136:864–871, 1964.
- [2] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, 1965.
- [3] J. P. Perdew and A. Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B*, 23:5043–5079, 1981.

- [4] D. C. Langreth and J. P. Perdew. Theory of nonuniform electronic systems. I. Analysis of the gradient approximation and a generalization that works. *Phys. Rev. B*, 21:5469–5493, 1980.
- [5] J. P. Perdew and Y. Wang. Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation. *Phys. Rev. B*, 33:8800–8802, 1986.
- [6] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77:3865–3868, 1996.
- [7] A. D. Becke. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.*, 98:5648–5652, 1993.
- [8] C. Lee, W. Yang, and R. G. Parr. Development of the Colic-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, 37:785–789, 1988.
- [9] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.*, 98:11623–11627, 1994.
- [10] T. Helgaker, P. Jorgensen, and J. Olsen. *Molecular electronic-structure theory*. Wiley, 2000.
- [11] M. Fuchs, Y. M. Niquet, X. Gonze, and K. Burke. Describing static correlation in bond dissociation by Kohn-Sham density functional theory. *J. Chem. Phys.*, 122:094116, 2005.
- [12] M. Grüning, O. V. Gritsenko, and E. J. Baerends. Exchange-correlation energy and potential as approximate functionals of occupied and virtual Kohn-Sham orbitals: Application to dissociating H₂. *J. Chem. Phys.*, 118:7183–7192, 2003.
- [13] P. Gori-Giorgi, M. Seidl, and G. Vignale. Density-functional theory for strongly interacting electrons. *Phys. Rev. Lett.*, 103:166402, 2009.
- [14] M. Seidl, P. Gori-Giorgi, and A. Savin. Strictly correlated electrons in density-functional theory: A general formulation with applications to spherical densities. *Phys. Rev. A*, 75:042511, 2007.
- [15] P. Gori-Giorgi and M. Seidl. Density functional theory for strongly-interacting electrons: perspectives for physics and chemistry. *Phys. Chem. Chem. Phys.*, 12:14405–14419, 2010.
- [16] F. Malet and P. Gori-Giorgi. Strong correlation in Kohn-Sham density functional theory. *Phys. Rev. Lett.*, 109:246402, 2012.
- [17] G. Monge. Mémoire sur la Théorie des Déblais et des Remblais. *Histoire Acad. Sciences, Paris*, 1781.
- [18] L. V. Kantorovich. On an effective method of solving certain classes of extremal problems. *Dokl. Akad. Nauk. USSR*, 28:212–215, 1940.
- [19] L. V. Kantorovich. On the translocation of masses. *Dokl. Akad. Nauk. USSR*, 37:199–201, 1942.
- [20] C. Villani. *Optimal transport: Old and new*. Springer, Heidelberg, 2009.
- [21] C. Cotar, G. Friesecke, and C. Klüppelberg. Density functional theory and optimal transportation with Coulomb cost. *Comm. Pure Appl. Math.*, 66:548–599, 2013.
- [22] G. Buttazzo, L. D. Pascale, and P. Gori-Giorgi. Optimal-transport formulation of electronic density-functional theory. *Phys. Rev. A*, 85:062502, 2012.
- [23] C. Cotar, G. Friesecke, and B. Pass. Infinite-body optimal transport with Coulomb cost. *arXiv:1307.6540*, 2013.

- [24] G. Friesecke, C. B. Mendl, B. Pass, C. Cotar, and C. Klüppelberg. N -density representability and the optimal transport limit of the Hohenberg-Kohn functional. *J. Chem. Phys.*, 139:164109, 2013.
- [25] C. B. Mendl and L. Lin. Towards the Kantorovich dual solution for strictly correlated electrons in atoms and molecules. *Phys. Rev. B*, 87:125106, 2013.
- [26] S. Vuckovic, L. O. Wagner, A. Mirtschink, and P. Gori-Giorgi. In preparation.
- [27] F. Malet, A. Mirtschink, K. J. H. Giesbertz, L. O. Wagner, and P. Gori-Giorgi. Exchange-correlation functionals from the strongly-interacting limit of DFT: Applications to model chemical systems. *arXiv:1401.7822*, 2014.
- [28] E. H. Lieb. Density functionals for Coulomb systems. *International Journal of Quantum Chemistry*, 24:243–277, 1983.
- [29] M. Seidl. Strong-interaction limit of density-functional theory. *Phys. Rev. A*, 60:4387–4395, 1999.
- [30] M. Seidl, J. P. Perdew, and Mel Levy. Strictly correlated electrons in density-functional theory. *Phys. Rev. A*, 59:51–54, 1999.
- [31] W. Gangbo and R. J. McCann. The geometry of optimal transportation. *Acta Math.*, 177:113–161, 1996.
- [32] M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof, and T. Østergaard Sørensen. Electron wavefunctions and densities for atoms. *Ann. Henri Poincaré*, 2:77–100, 2001.
- [33] PHG, <http://lsec.cc.ac.cn/phg>.
- [34] MOSEK, <http://www.mosek.com>.
- [35] A. J. Cohen, P. Mori-Sánchez, and W. Yang. Challenges for density functional theory. *Chem. Rev.*, 112:289–320, 2012.